DOCUMENT RESUME

ED 392 838 TM 024 614

AUTHOR Kvale, Steinar
TITLE Evaluation as Construction of Knowledge.
PUB DATE Apr 95
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Constructivism (Learning); Cross Cultural Studies; Culture; *Evaluation Methods; Foreign Countries; Intellectual Disciplines; *Knowledge Level; Metacognition; *Student Evaluation; *Test Use
IDENTIFIERS Discourse; Knowledge Acquisition; *Metanarration; *Postmodernism

ABSTRACT

Evaluation is discussed as knowledge construction. Examinations are considered as tests of knowledge or as tests of students and then as selection tools or as the construction of knowledge in themselves. Examinations involve the power to define what is common knowledge in a discipline and in a culture, and the examination discourse contributes to the construction of this valid knowledge. The postmodern perspective on knowledge is characterized by a disbelief in the metanarratives of legitimation. Knowledge is not fixed, but constructed, and a postmodernist perspective leads to the interpretation of knowledge as social construction. Some of the knowledge tensions in a postmodern condition are applied to evaluation, focusing on the tension between universal commensurability and the narratives of local cultures in an era when universal metanarratives no longer appear as the foundations of valid knowledge. The future of evaluation may rest on research that crosses cultural and disciplinary boundaries. (Contains 2 figures and 31 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 392 838

American Educational Research Association - Annual Meeting
San Francisco, April 18 - 22, 1995

EVALUATION AS CONSTRUCTION OF KNOWLEDGE

Steinar Kvale
University of Aarhus
Denmark

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
STEINAR KVALE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Introduction

1. **A historical introduction**

2. **Evaluation of students or of knowledge?**

3. **Examination as knowledge construction**
3.1 Postmodern construction of knowledge
3.2 Evaluation as knowledge construction
3.3 Evaluation as thematic knowledge constitution and censorship

4. **Evaluation as structural knowledge constitution**
4.1 Discursive knowledge
4.2 Differentiation of knowledge
4.3 Fragmentation of knowledge
4.4 Knowledge as privilege
4.5 Commodification of knowledge
4.6 Castration of knowledge

5. **Knowledge and evaluation in a postmodern condition**
5.1 Loss of legitimacy
5.2 Local narratives
5.3 Capitalization of knowledge
5.4 Commensurability

Concluding perspectives

References

TM 024 614

BEST COPY AVAILABLE

Be commensurable, or disappear (Lyotard)

I will here adress evaluation as knowledge construction. First I shall treat the topic of exams as tests of students or as tests of knowledge - of exams as selection of students versus exams as construction of knowledge. Then I will introduce a postmodern perspective on knowledge as social construction, and outline how the examination discourse contributes to the construction of the valid knowledge of a discipline. Finally some of the tensions of knowledge in a postmodern condition will be applied to evaluation, in particular a tension of universal commensurability and the narratives of local cultures.

1. A historical introduction

Evaluation is a new term in education and has in last decades been replacing traditional, and often negatively laden, terms as examinations and grades. Before turning to current evaluation, some aspects of the history of examinations shall be briefly mentioned.

The prehistory of examinations goes back to the Han dynasty in China 200 years before Christ. The comprehensive civil examination system was developed by the Emperor T'ai-tsung in the seventh century after Christ. The exams were based on essays treating literature and the thought of Konfutse, they were strictly controlled, and only a few of every hundred candidates passed through the provincial examinations to the equally hard final metropolitan palace exam, which was supervised by the Emperor himself. While being strenous - a recent history thus bears the title "China's Examination Hell" (Miyazaki, 1976) - the examinations were fair, and immeasurably progressive at the time. The exams were in principle open to all qualified applicants. The main function of the bureaucratic examination system was to recruit loyal civil cervants to the emperor; T'ai-tsung thus remarked when observing a procession of the candidates who had passed the final palace exam: "The heros of the empire are all in my pocket!" (Mizyaki, 1976, p. 113).

The introduction of the civil service examinations in the seventh centruy broke the rule of the hereditary aristocracy, and the legitimacy of the civil cervants recruited by the exams served to keep the military officers subordinate to civil rule. It is thus possible, as argued by Mizyaki, that the recruitment of loyal civil servants through the examination bureaucracy served to keep China a united empire at a period when Europe was caught in continual feudal wars of provinces and nations.

The civil service examinations became too rigid and they were abolished at the turn of this century, with examinations becoming a continually debated topic. Thus Mao Zse Dong was rather critical of exams: "The present examination system appears more suitable for enemies than for the people; it is like an ambush, because the questions are unpredictable, special, and still held in the tradition of the essays in eight parts...The examination system fights the pupils like enemies. It is desctructive and should be prevented (quoted after Ch'en, 1974, p. 94).

In Europe examinations were introduced at the universities at the end of the middle ages. They were mainly in the form of public oral disputes, which could be rather harsh (see Kvale,1972). The founder of the university at Paris, Robert de Sorbon, thus in the thirteenth century compared the university exams with the Christian day of the last judgement. He found as the main difference that the judges over heaven and hell were by far milder that the judges at the university examinations. The exams of the period were aggressive; thus at the University of Cambridge the candidates had to swear before the examination that they would never carry revenge on their examiners. The examination rules at the University of Heidelberg from 1501 contained a passage forbidding the students to carry long knives at the examination.

In the modern era the ritual oral examinations were replaced by the formal bureauchratic exams. The grades, introduced in the European schools by the Jesuits in the sixteenth century, became a central mean of control and selection. Grades represent a bureaucratic stroke of genius, the one-dimensional grading scale makes possible a simple formal selection of pupils for further educational privileges. The grades appear objective and legitimate, they are based on the pupils own achievements, with every pupil having a formally equal chance of getting to the top of the educational hierarchy.

While it is often popular among students to bash at the formalistic bureaucracy of exams and grades, it represents a formal legal system, with access to higher education through own merits, rather than family position and income, political correctness and connections through the back door. After the selection to higher education through political virtues during the Cultural Revolution in China, the selection through exams and grades were by many students experienced as a liberating force, as described by Jung Chang in her autobiography "Three Swans" (1992; see also Thoegersen, 1990).

2. Evaluation of students or of knowledge?

Evaluation means according to the dictionary: to ascertain the value of something; evaluation is used for determining the price of a commodity. In education, evaluation pertains to the evaluation of students' knowledge. This formulation is, however, ambiguous, is it the students who are evaluated on the basis of the knowledge they document, or is it the knowledge presented by the students which is evaluated?

The dual topic of evaluation shall be discussed here in relation to a common form of examination at Scandianavian universities where an essay is evaluated by two examiners. One is a teacher at the institution, and the other is an external examiner, often called "censor", who comes from another university or from the local professional community. The evaluation of the paper is sometimes followed by an oral examination of, or a more open dialogue with, the student about the paper.

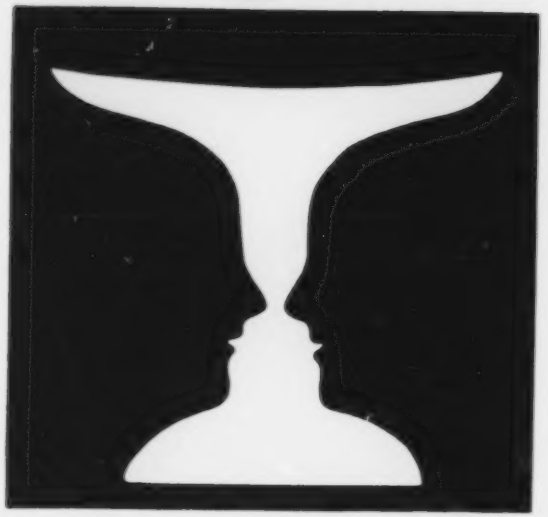

Figure 1. A figure-ground understanding of evaluation of student or of knowledge

Figure 1 depicts the two alternative ways of conceiving the subject matter of exams - as an evaluation of students and as an evaluation of knowledge. The faces depict the two examinators of an examination committee, or the student and the examiner, and the vase inbetween may be seen as the knowledge tested. The figure emphasises the interdependence of students and knowledge, and also the possibility to abstract and focus upon one of the two - students or knowledge.

Debates and research on examinations have generally focused upon the students evaluated. I shall here discuss examinations as a test of knowledge. Through evaluation of the students' presentation of their knowledge also the common knowledge of a discipline is evaluated. Examinations involve the power to define what is the common valid knowledge in a discipline and in a culture. Evaluation ascertains the value of the knowledge presented, as well the form of evaluation itself promotes certain values.

Function:	*Subject matter:*	
	Students	Knowledge
Selection	1	3
Transformation	2	4

Figure 2. Subject matter and functions of educational evaluation

In Figure 2 two main functions of evaluation - selection and transformation- are posited for the subject matter as students or knowledge. The outcome of evaluation - usually in the forms of grades - is used to select pupils for further educational privileges (1). With evaluation as the de facto goal of learning, the forms of evaluation come to influence and transform the students' learning (2). The same exams also involve a selection of what knowledge is important enough to be included by evaluation (3). And, to be discussed later, the evaluation may transform the knowledge evaluated (4). We then arrive at two alternative understandings of the function of evaluation - as selection and transformation - and of two alternative conceptions of the subject matter of evaluation - as students or as knowledge.

When I first started studying examinations in the late 60s in Oslo, my focus was on square 1 in the figure. This took place by a statistical study of the reliability of grading at examinations in psychology at the University of Oslo in Norway (Engvik, Havik & Kvale, 1970). This lead to an interest in the effects of evaluation upon students, moving to square 2 in the figure, and drawing in scattered examples from empirical studies and some implications of general principles of learning and motivation (Kvale 1970). Later again I conducted interview studies among Danish high school students on the effects of grading upon learning and social relations in school (Kvale, 1980).

The knowledge aspect of exams in square 3 and 4 mainly came to my attention as an oppressive censorship of meaning (1972). Later I came to focus on the knowledge evaluated, changing my earlier critique to regard examinations as necessary aspects of the maintainance and construction of the knowledge of a discipline (1990). In the present context the important issue of knowledge excertion over students through examinations will not be treated. The focus shall here be on the more neglected power excertion through examinations in defining the knowledge of a discipline.

Before turning to knowledge construction by examinations, a brief overview of research on examinations as selection and transformation of students shall be given. I will estimate that about 90% of research on examinations is to be found in square 1 in Figure 2, concerning selection of students. Here are far more studies of reliability than of validity, and again less on the social selection aspects. Research on evaluation as selection of students show that grades generally are not very reliable or valid measures of students, they are not effective predictors of student performance. The grades do, however, allow a simple meritocratic and bureaucratic selection to privileges.

Of the remaining 10% examination research, a majority is found in square 2 on the influence of exams in transforming students. This has mainly concerned feedback on learning and exams as an motivational device, and little on the influence of exams on attitude formation and personality development. When grades are the key to advancement in the educational hierarchy, the grades may become the primary goal of learning. There occurs a means-ends reversal, where the grades as means of measuring learning become the purpose of learning, the measurement of learning becomes the meaning of learning. The common forms of exams and grading are not, according to psychological principles of learning, efficient promoters of the curricular goals of learning and understanding. Here appears a contradiction between the testing and the educative functions of exams: as discussed by Wiggins (1993) what is educationallty vital is inherently at odds with efficient testing and unambiguous test items. Grades appear more important, although here little research has been conducted, when it comes to disciplining students in school and socialising them to the wage labour of a capitalist economical system (Becker, Geer & Hughes, 1968; Bowles & Gintis, 1976.)

Concluding this brief overview of research on students as the topic of evaluation, we may state that while grades are not very efficient predictors of future success, they may function as a legitimate bureaucratic selection procedure. And while the common forms of evaluation are not very efficient in promoting learning, grading may be important means of discipling and socializing students.

We shall now turn to the alternative conception of evaluation as a test of knowledge, outlined earlier (Kvale, 1990, 1993a) and to be discussed here from a postmodern perspective upon knowledge. Whereas the influence of tests upon the pupils and students tested have been discussed (e.g. Hansson, 1993), the impact of tests and exams upon the knowledge tested has been little investigated. There has been some discussion on the selection of the knowledge tested with regard to content validity in relation to curricular goals. The contributions of examinations to the social construction of the knowledge of a discipline has, however, hardly been treated. One exception is Delandshare & Petrosky's (1994) Foucault inspired analysis of teacher assessment as knowledge creation with a discussion of the function of evaluation in defining what constitutes knowledge.

3. Examinations as knowledge construction

3.1 Postmodern construction of knowledge

The postmodern condition is characterized by a disbelief in metanarratives of legitimation, such as Enlightment through knowledge or the emancipation of the workers (Lyotard, 1984). With a delegitimation of global systems of thought, there is no foundation to secure a universal and objective reality. The dichotomy of universal social laws and unique individual selves is replaced by the interaction of local networks, the self is decentered to become an ensemble of relations.

The focus is on the local context and on the social and linguistic construction of a perspectival reality where knowledge is validated through practice (Kvale, 1992).

The belief in an objective reality has been at the basis of a modernist understanding of truth and validity. The philosophy of Enlightment was a reaction against the religious dogma of the so-called dark ages. The medieval belief in one true and almighty God, for all people, from eternity to eternity, was in the modern era replaced by the belief in one true and objective reality, stable and universal. In a postmodern era the foundations of true and valid knowledge in an absolute God, or an objective reality, have dissolved. The conception of knowledge as a mirror of reality is replaced by knowledge as a social construction of reality. There is a focus upon interpretation and negotiation of the meaning of the lived world. Knowledge is not a matter of interaction with a nonhuman reality, but of communication between persons; as argued by Rorty (1979) the conversation becomes the ultimate context within which knowledge is to be understood. Truth is constituted through dialogue; valid knowledge claims emerge as conflicting interpretations and action possibilities are discussed and negotiated among the members of a community.

I shall here discuss some implication of a postmodern perspective upon examinations. A conception of examinations as knowledge construction is not restriced to a postmodern understanding, but it becomes prominent with a postmodern decentering of the individual and the understanding of knowledge as constructed through discourse. There is a *decentering* of the individual student and teacher of the examination to a focus on their positions in a network of interchange and transmission of knowledge. It is less the individual subject who contains and uses knowledge than knowledge which posesses and uses the subjects. Knowledge is not fixed entitities transmitted from one subject to another, knowledge is linguistically and socially *constructed* through interaction. There are multiple *discourses* involved in the creation of the knowledge in a discipline: scientific research, public debate, teaching, construction of curricula and pensa, and to be discussed here - the examination discourse.

3.2 Evaluation as knowledge construction

The examination system controls the content and the form form of knowledge passed from one generation to the next. The examination controls the relation between traditional and new knowledge, between universal and local knowledge. A discipline's knowledge is not merely transmitted in teaching and evaluated at examinations; the examination itself contributes to the construction of the valid knowledge. The very operation of examining provides an operational definition of the official knowledge of the discipline.

The examinations define for the students what knowledge that is worthy of acquisition and mastery. The examination is an authoritative statement about what knowledge is outside of a discipline; for example "This may be a brilliant philosophical essay, but it falls outside of the field of a scientific psychology, and is thus failed". Within a knowledge conception the students evaluated are "bracketed", they are mere carriers of the knowledge to be offically evaluated. The exam candidates are the yearly lambs, some of whom have to be sacrificed at the examination table, in order for the community to maintain the necessary boundaries for its common valid knowledge.

The question of what is a valid examination involves different answers with different conceptions of topic and purpose of examinations. We may here return to the four squares of figure 2 above. With the purpose of evaluation as selection of students a predictive validity for student's future behavior is involved (1). With the purpose of evaluation as transforming students we may talk of a pragmatic validity, the evaluation serves to influence the students knowledge and behavior (2). With evaluation as a selection of knowledge a content validity pertains; that is whether the examinations select and grade the knowledge depicted by the official goals and curricula of a discipline (3). And with evaluation as a social construction of knowledge we may talk of a communicative validity, requiring a dialogue and a decision of what is the valid knowledge of a discipline (4) (see Kvale,1995, for a discussion of pragmatic and communicative validity).

We shall now treat two kinds of transformative effects of examinations upon knowledge - thematic and structural.

3.3 Evaluation as thematic knowledge constitution and censorship

A *thematic* influence of examinations refers to defining the content, what knowledge, what theories and viewpoints, are acceptable in a discipline. This is an issue sometimes discussed as censorship of knowledge. The evaluation of the worth of a student's examination paper rests upon a conception of what is valid knowledge in a discipline. It involves taking a position on what is true and what is false knowledge as well as on what knowledge is inside and what is outside of a discipline. Taking a clear position on what is true knowledge may, in a liberal society, be rejected as being an outdated censorship of meaning stemming from oppressive religious and political regimes.

In the Scandinavian and German languages "censur" has long been used as an official term for examinations and for grading. It is currently being replaced by the term "evaluation", which however, as indicated by the word meaning, still implies a position taking on values. Today reports of censorship of knowledge tend to be rejected as extreme and atypical, censorship is regarded as a misuse of the examination. It is still possible, however, that the freedom of thought expounded in lectures may be undermined by institutional censorship of meaning at the examination, thus upholding an "illusion of free thought".

A few decades ago, within the discipline of psychology in Scandinavia, controversies arose over whether psychoanalytical, critical and Marxist thought could be included in the official knowledge of a scientific psychology. In given periods and university departments, students who believed the knowledge claims of Freud, Habermas, and Marx were relevant for the science of psychology may have refrained from mentioning them in their examination papers. Those who did mention such ideas risked to be failed.

In such cases there may be accusations of a censorship of meaning - between students and examiners, or, within an examination committee, between different examiners evaluating the same paper. These disputes about the truth and the value of the knowledge evaluated are some-

times accompanied by strong emotional and personal conflicts, between the examiners, or between examiners and student. The basic issue may then be less the particular pass/fail to be assigned to the student, but what is experienced as an attempt to invalidate one's postion on what is the true valid knowledge of the discipline. A few years later the previously outlawed theories may become incorporated into the mainstream of the discipline. Thus in Scandinavia, psychoanalysis, the Frankfurt school of critical theory, and Marxism, have since 1968 marched the long way through the examination systems, and they are now in general considered legitimate ways of understanding in psychology.

At this point I would like to present the argument that examinations are not merely an oppressive censorship of knowledge, and argue for a discursive understanding of examinations as a necessary institution where the valid knowledge of a discipline is negotiated and constructed. The examiners decide what knowledge shall be selected as acceptable for meeting the standards of a discipline. Thus examinations implicitly define the very standards and values of a discipline, which I shall refer to as *meta-evaluation*.

In normal stable periods the meta-evaluative aspects of an examination is hardly visible. In times of radical paradigmatic changes in knowledge, when the traditional values of a discipline are challenged, there is a necessity to redefine the limits of the discipline, to re-establish its core concepts, methods and data. During periods where a discipline simultaneously contains incommensurable frames of reference, a common set of standards needs to be established anew, in order to obtain consensually valid evaluations of a student's knowledge. In such periods of upheaval, where a number of knowledge claims within a discipline are in conflict with each other, the censoring, meta-evaluative, aspects of evaluation becomes visible.

In normal periods of a discipline examinations serve mainly as a routine maintainance of knowledge, confirming and certifying the validity of the knowledge brought forth by the students. At lower levels of the educational system the maintainance function of evaluation dominates, whereas at higher levels of the educational hierarchy conflicts of the knowledge construction may be more visible. By university examinations, by ph.d dissertations and by faculty appointments, the worth and validity of the more complex and new knowledge evaluated may become openly contested.

The present position is that the evaluation system is a nexus, a point of control, where concrete decisions are made defining what knowledge is acceptable for inclusion in the curriculum. The knowledge of a discipline is not constituted solely by examinations; scientific research, and also public debate, create the communally accepted knowledge. Within the educational system curriculum plans indicate the official knowledge goals and the teaching transmits the official knowledge of a discipline. The generally vague definitions of goals of learning and the lack of explicit criteria of evaluation, however, leave large margins for interpretation by the examiners: through their concrete evaluative decisions, they operationally define what is the valid knowledge of a discipline.

Production and verification of knowledge takes place in scientific research, in professional practice and in public debates. The earth will not become flat, nor stand still if an examination committee should decide so. It is, however, the examiners who decide the locally valid and legitimate knowledge, demanding an understanding of science which the students will have to pledge allegiance to before they are admitted through the examination gate into the profession.

4. Evaluation as structural knowledge constitution

Examinations do not only contribute to the construction of the valid content of a discipline, they also influence the very structure of the knowledge evaluated. A *structural* influence of the exams refers to the more indirect and less visible influence of the forms of evaluation upon knowledge, such as of different examination types - oral, essay and multiple choice examinations - and of different forms of grading, as by ranking or by pass/fail. The structural effects of the forms of examination upon knowledge tend to be less visible, and have been less debated, than the issue of censorship of knowledge content. I shall here outline six implications of evaluation structures for the knowledge evaluated: knowledge as discourse, knowledge differentiation and fragmentation, knowledge as privilege and commodity, and finally castration of knowledge.

4.1 A discursive knowledge conception

In a discursive approach knowledge is seen as a social construction, where valid knowledge emerges as conflicting knowledge claims are argued and accepted in a dialogue with a resulting consensus. Evaluation involves a negotiation of the true knowledge of a discipline. A discursive conception of knowledge was involved in Socrates' dialogues as well as in the harsh medieval, and the milder modern, public defences of the knowledge claims of a dissertation.

In the Scandinavian examination committees the validity and the value of student's knowledge claims are argued in a committee of two examiners, and in some cases by the student in an oral defense. The examination committee serves as representative of the interpretative community, assessing the validity of a discipline's knowledge as presented by the students at the examinations. The committees are faced with the task of maintaining and renewing the traditional body of knowledge, deciding what new lines of thought are to be included in the discipline and what is to be rejected. In this view the work of an examination committee approximates the hermeneutic interpretation of texts in the humanities. The task is through a dialogue to arrive at a consensus of the meaning of the text and of its quality within the interpretative tradition of a discipline.

4.2 Differentiation of knowledge

From a knowledge constitution viewpoint, a finely differentiated grading scale may contribute to a preciseness of knowledge. Thus if the task of an examination committee is merely to decide whether an examination essay is to pass or fail, the conception of knowledge at the basis of the decision need not be particularly precise. If the task is, however, to arrive at a consensus about the position of an essay on a finely differentiated scale of quality, this requires a precise and intersubjective concept of valid knowledge.

Compared to other disciplines, law has often a rather strict examination procedures and a finely differentiated grading scale. From a knowledge constitution viewpoint the strict evaluative dif-

ferentiation of legal knowledge involves a coercion towards establishing precise and consensual interpretations of the legal texts. And with regard to effects upon students the strict differentiation by evaluation may be seen as socialization of students, with a particulary strong necessity of establishing discipline, loyalty, rank order and respect for authority within the legal profession.

4.3. Fragmentation of knowledge

The evaluative objectification of knowledge which may generally be found with a bureauchratic examinations system is particularly strong by multiple choice tests. Here the students are to choose between pregiven answer alternatives to the examination questions. The response alternatives are either true or false, thereby screening out the subleties and ambiguities of complex forms of knowledge. A contexual understanding of knowledge is excluded, what matters are the fragmented statements about objects of information to be marked as unequivocally true or false. In contrast to essays, the multiple choice tests do not place the emphasis on the production of coherent ideas and logical argumentation, but on the "recognition of small pieces of mostly factual, and often trivial information. Intolerant of ideosyncracy and individuality, it carried the latent message that the goal of learning [is] the identification of the one 'wholly right' answer as defined by seemingly impersonal, 'objective, yet often quite arbitrary authority" (Samelson, 1987, p. 123-4).

Despite empirical investigations indicating that multiple choice tests may promote a fragmented and superficial conception of knowledge, emphasising memory of facts over general comprehension and critical evaluation (see Balch,1964; Fredriksen,1984; Madaus,1990) these "side effects" of the measurement of learning have, in spite of their contrast to curricular goals, seldom been recognised in the educational literature.

In the multiple choice examinations knowledge is constructed as a matter of facts and rules for combining facts. This is in accordance with an empiristic philosophy of knowledge, as well as with the late artificial intelligence tradition of simulating human understanding with computers. The empiristic conception of knowledge has come under critique from philosophers and computer scientists for neglecting the importance of contextual knowledge, ambiguous knowledge and the implicit background knowledge necessary for understanding facts and rules the (Dreyfus, 1993).

4.4 Knowledge as privilege

Through examinations the knowledge of a discipline becomes a privilege, a private property which belongs to those who have passed the examinations. They obtain the certified right of using the knowledge in their professional practice as medical doctors, laywers etc. The examinations secure a monopoly on the use of, and getting reimbursed for applying, the discipline's knowledge. It may be forbidden by law for those who have not obtained the examinal privilege, and membership of the guild, to use this knowledge, for example in medicine.

Examinations are a certification of knowledge as privately owned, or rather they give access to membership of a guild who owns the rights to use the discipline's knowledge. This certified

knowledge creates a monopoly, rigidly and legalistically declaring who can operate as a professional in a given field. Punishment is mated out on those who use the professional knowledge without the ownership of a diploma or certificate won through the examination process. The examinations serve as boundary markers between different professional disciplines, defining what is inside and what is outside a discipline of knowledge.

The examination system may contribute to a fragmentation of knowledge by counteracting interdisciplinary crossings of disciplinary boundaries. The present discussion has thus attempted to understand examinations from the viewpoints of test psychology, from educational learning, and from sociology and epistemology of knowledge. These viewpoints, however, belong to different disciplines, each with their separate examinations. The resulting compartmentalisation of knowledge about exams may be observed in educational literature on exams as well as at educational conferences.

Karl Marx has in his critique of Hegel's philosophy of law suggested an understanding of examinations as a knowledge privilege. Marx describes an examination as a legal recognition that official knowledge is a privilege; the examination is the objective link between the official position and the individual, which together with the princely grace gives access to higher public positions: "The examination is nothing but the bureaucratic baptism of knowledge, the official recognition of the transubstantiation of profane knowledge into sacred knowledge" (Marx, 1961, p. 253). Paraphrasing Marx we may say that today the role of the examination has become a bureauchratic technologization of knowledge.

4.5 Commodification of knowledge

Evaluation in grades makes knowledge something to be measured and quantified, everything can be compared with everything; an A in mathematics is exactly identical with an A in religion. Knowledge becomes a commodity to be acquired and exchanged, containing the contradictions of a commodity as analysed by Marx. The use values of knowledge are qualitatively diverse, such as knowledge of biology used for increasing agricultural production, for ecological interventions, or for esthetical enjoyment through better discernment of the wild life. The exchange value of knowledge is quantifiable and commensurable, it is made out of what other commodities knowledge can be exchanged for - in economic life for the general exhange equivalent money, and in the educational system for grades as the exchange equivalent.

An instrumental commodity understanding of learning was expressed in this way by a Danish high school pupil: "My interests have moved away from what takes place in school. I go here with the explicit purpose of getting as good an examination as possible, with the least possible effort" (Kvale, 1980). The students acquire an understanding of knowledge as commodity through the grading of the knowledge they present at examinations. Evaluation comes to mean ascertaining the price of the knowledge delivered by the students. The primary value of the knowledge acquired is not what it can be used for, but what it can be exchanged for, to which privileged education the grades give access to. A strong focus on evaluation in education, with high stakes for students and teachers in the outcome of evaluation, may actually lead to a devalu-

5.2 Local narratives

With the collapse in the postmodern condition of universal meta-narratives as foundations of valid knowledge, the local narratives come into prominence. The particular, hetereogenous and changing language games replace a global horizon of meaning. The interaction of the participants of the local community and of their local knowledge become important in their own right. The indigenous local knowledge enters into the cultural interchange of a pluralistic multicultural world.

We see today a search for local compensations for general loss of legitimacy. There may be among the youth a search for local masters, for authorities which the students choose themselves, as a compensation for a school with little legitimacy and the loss of belief in the general values of society. On the administrative level there is a movement towards decentralisation, which with a loss of the authority of the modern state and of traditional knowledge, may function as a compensatory form of legitimation (Weiler, 1990).

If universities are to retain their classical tasks of transmission and creation of culture they need to retain control of the evaluation system. The communal construction of valid knowledge then remains to be worked out and negotiated in relation to the local culture, rather than by new multinational educational testing agencies. A discursive form of evaluation, discussed above, can make it possible to maintain a dialogue of traditional and new knowledge, of local and of general knowledge.

5.3 Capitalization of knowledge

In scientific research in a postmodern condition "the goal is no longer truth, but performativity, that is the best possible input/output equation". The State and the companies abandons the idealist and humanist narratives of legitimation: "in the discourse of today's financial backers of research, the only credible goal is power. Scientists,'technicians, and instruments are purchased not to find truth, but to augment power" (Lyotard, 1984, p. 46).

Within education, goals as fostering a love for true knowledge, and as promoting the moral and spiritual values of a society, may still be mentioned in ceremonial speeches, but they rarely occur in the official language educational reforms. Here the language game of performatity reigns: in order to estimate the yields of the investments in education it is necessary to have to have clear standard goals for the outcomes of education, with measurable outcomes of the quality of education. The economic disourse of performativity and accountability replaces in education the cultural discourse of knowledge and truth. The current conceptions of education as investment in human capital a may serve as a new form of legitimation; after the grand legitimation narratives of the Enlightment - emancipation and liberation through education - are breaking down, the performativity of capital profits becomes the ultimate frame of reference for legitimacy of education.

Today the contributions of the World Bank and Unesco to the development of national educational systems is based on assumptions of universal applicability of Western ideas of knowledge and educational planning (cfr. Weiler, 1989). Education as investment in human capital is a main aspect of official educational policy in many countries. The local indigenous culture may remain as harmless entertainment in folkloristic reservations and in the new Disneylands.

The current emphasis on international comparisons of educational outcomes requires a commensurability of the evaluation results. International comparisons may be valuable feedback on the level of education in a country, but if high stakes become invested in international comparisons, the message for the indigenous knowledge and values of a country is - become commensurable or disappear. The emphasis upon international comparisons may also in some countries have internal uses. The alleged educational gap to other countries may serve as a bogus gap, after the design of the international arms race in the 80s. Here the evocation of increased external military threats may also have served to enhance internal countrol in the own country as well as to promote its weapon industry. The alleged knowledge gap in education may also be used to increase internal control through standardisation and technologisation of the educational system and of the knowledge transmitted. The demands of international commensurability of knowledge way serve as a weapon against the knowledge contained in the discourses and narratives of the local communities.

5.4 Commensurability

With a dominance of a performativity principle in the postmodern condition, social interaction is managed by input/output matrices which imply commensurable elements. The legitimation of power is based on its optimising of the system's performance-efficiency, which entails a certain element of terror:..."be operative (that is commensurable) or disappear" (Lyotard, 1984, p. xxiv).

With a globalization of economy a heterogenous plurality of cultures is need no longer be threatening to the institutions of power; today major power and control takes place through international currency exhanges. A universal culture of commodities takes over, with money as the one common value across, and above, the postmodern pluralism of cultural values. In the 16th century Galileo put forth the universal demand: measure what is measurable, and make measurable what is not. At about the same time the Jesuits started measuring and comparing their students' knowledge and on a grade scale.

Today the common currency of the educational system - the grades - ensure control and loyalty more effectively than the knowledge and values directly taught. The fostering of true knowledge and a common body of values is becoming obsolete. The Western missionary command has lost its power, it is no longer: go out and make all mankind my disciples. The modern dictum has become: go out and make all mankind measurable - and potential customers of commodities.

The demand for commensurable forms of educational evaluation, within the classroom, within the nation, and lately internationally, looms large in education. The OECD has initiated a major project for improving the collection and dissemination of internationally comparable indicators of education systems (INES). In the European Community a system has been developed for measuring and comparing the student's learning achievements, and transferring them from one

ation of the content, the knowledge and values taught; what matters is merely the numerical outcome of the evaluation. With such an emptying devaluation of the knowledge content of examinations, a formal censorship of the meanings expressed may become less important.

4.6 Castration of knowledge

Evaluation of knowledge may extract the force and potency of the knowledge. A strong focus on grading, with high stakes for the students in the outcome of grading, may further an evaluative erosion of knowledge. The Jesuits introduced the grading system in European schools in the sixteenth century, possibly through the experiences of Jesuit missionaries in China. With the aim of differentiating clearly among the pupils they introduced a scale of 1 to 6. The grades were mainly for internal use in the classroom; the pupils were ranked according to their grades and received honorary positions such as rank of seat order. The Jesuits' main and explicit purpose for the use of grades was to promote discipline, competition and diligence. In their treatise on education, published in 1599, they emphasised competition as a strong force in promoting diligence (#39, Ratio Studiorum, 1887). Classroom discipline was obtained by the hope of honor and the fear of shame trough competition rather than by corporal punishment.

The French sociologist Durkheim (1977) has postulated an intended effect of the Jesuits' grading upon the knowledge acquired by the pupils. After the Rennaissance in Europe the learning of the classical Greek and Latin texts was part of the general culture which the pupils were supposed to know. The Jesuits, however, did not want their pupils to acquire any deeper and integrated knowledge of these pagan texts of antiquity. They therefore sought to lead the pupils attention away from the content of the dangerous texts and towards their form; the language, the grammar and style. A strong competition for grades was one of the pedagogical means for securing a superficial knowledge of the classical texts. The effects of the Jesuit schooling upon their pupils knowledge of the Greek and Latin culture was criticised by contemporaries, one critique thus described the Jesuit mode of education as a "castration of books".

5. Knowledge and evaluation in a postmodern condition

I have now discussed evaluation as knowledge construction and pointed out how different forms of evaluation reflect and further different conceptions of knowledge. I shall conclude by asking more specifically to the nature of knowledge in a postmodern world, and discuss the implications for evaluation. I shall take the book "Knowledge in the postmodern condition" (1984) by the French philosopher Lyotard as the point of departure. This was a report commisioned by the goverment of Quebec; as material for university reforms it outlined the conditions of knowledge in a postmodern era. Here I shall draw in some implications of analysis of postmodern knowledge for the evaluation of knowledge with regard to the tension between universalisation of knowledge and the rootedness of knowledge in the narratives of the local cultures.

With the breakdown of the modern metadiscourses for legitimation of knowledge, two contrasting sources of legitimation appear in a postmodern era - economical peformativity and the local interaction and narratives. On the one hand there is a recourse towards legitimation through performance-efficiency with an emphasis on universal commensurability. On the other hand there is

"the knowledge that legitimation can only spring from their own linguistic practice and communicational interaction" (Lyotard, 1984, p. 41). I shall here speculate on implications for evaluation of a dual legitimation of knowledge through general performativity and through local narratives.

5.1 Loss of legitimacy

Legitimacy involves the questions of what is valid, legal, and just. In a postmodern culture there is a pervasive disbelief in the grand systems of thought which have provided legitimation of knowledge. The belief in the such metanarratives as religion, the Enlightment belief in emancipation through knowledge, and the belief in a socialist utopia dissolve in a postmodern age.

With a loss in the legitimacy of the knowledge taught, we see in education a move from discussions of the truth and value of the knowledge taught to the formal methods of transmitting and evaluating the knowledge. Debates on censorship of meaning may be replaced by debates on how many steps should there be on the grading scale. Current demands for a reduction of exams, and abolishment of grading, with advocation of more open, individualised and interactive forms of evaluation, may be seen as reactions to a pervasive loss of legitimacy of traditional knowledge. If the knowledge learned is without an objective value, why evaluate it?

On the other hand, the standardised and technological evaluation procedures, such as multiple choice tests and grading along the normal curve, may serve as attempts to uphold the crumbling knowledge content. The appearance of objective measurement may provide a compensatory legitimation, in that the very focus on the evaluation techniques may serve to uphold a belief in the objectivity of the knowledge measured. What counts is what can be counted.

Common to the two positions of reducing or strengthening educational evaluation, is the focus on the forms of evaluation, rather than upon the content evaluated. With a loss of legitimacy of the knowledge evaluated, a censorship protecting true knowledge against false knowledge claims recedes. There is less a discussion of content, of the truth of the knowledge claims, than of the structural aspects of exams, such as examination forms and grades. This is here interpreted as an indication of problems with the justification of the knowledge evaluated: if the goals of learning were of primary importance, the methods to measure the knowledge would be secondary to the truth value of the knowledge.

Weiler (1990) has pointed to the phenomenal rise of system evaluation as in part due to its capacity of confirming scientific rationality and respectability upon the policy process, evaluation serving as a "compensatory legitimation". Applied to educational evaluation this implies that in face of a postmodern confusion and loss of authoritative values and valid knowledge, the very focus on the concepts and techniques of evaluation may serve to uphold a belief in the existence of educational values and valid knowledge. At the same time the strong focus on evaluation may, as discussed in relation to commodification of knowledge, lead to a devaluation, or evaluative erosion, of the knowledge and values to be learned.

institution to another. The quality of work is then represented on a new common grading scale from 1-7. We may speculate that with the local legitimacies of the national grading systems the introduction of a common currency in the educational system will encounter more resistance than the introduction of the common monetary currency ECU.

Cross-national comparisons of knowledge are today based upon multiple choice tests. For the sake of clarification we shall here analytically distinguish between three realms of knowledge - facts, meanings and values.

Multiple choice tests are well suited for testing knowledge of *facts and rules* combining the facts. For this kind of knowledge, which prevail in fields as natural science, medicine and mathematics, the standardised multiple choice tests in different countries may be approximately identical and make possible comparisons between knowledge in different countries.

Essay tests and oral examinations are better suited for evaluating understanding and interpretation of the *meaning* of literary texts in the humanities. The subject matter is here of more national nature than in natural sciences, making comparable examinations of national literature, of law and of theology across nations difficult. Where the very subject matter is crossnational, as much of history, crossnational evaluation may be especially hard to establish. The presentations of a war in the history books of the countries involved tend to differ. It may be difficult to obtain crossnational agreement on interpretations of causes, and also on facts of the wars taught about in the history books. Constructing a crossnational multiple choice test in recent history, with true and false answers about the facts and causes of major conflicts, may be a hasardous undertaking.

When it comes to evaluation of *values*, it makes little sense to test students' acquisition of moral and spiritual values through true/false answers on multiple choice tests. Here essays and oral discourse may give better indications of a students internalisation of the values promoted by education. With different fundamental values in different countries a cross-national evaluation of values appears difficult.

In conclusion, crossnational comparisons of knowledge can reasonably be made for realms of knowledge dominated by facts and rules. For knowledge involving interpretation of meaning and evaluation of values, crossnational comparisons are difficult; and when issues of national identities and values are involved they are hardly feasible. Cross-national comparisons in these realms of knowledge would involve a technological simplification of knowledge to facts and rules, and would require eliminating national differences of cultural interpretations and values. A universally valid test of human values appears futile.

We may today see a new Western intellectual colonisation, not by missionairies supported by soldiers, but by educational testers supported by bankers. The World Bank today requires loan seeking countries to adhere to the Western individual conception of Human Rights - No Human Rights, no money. Similarly the World Bank and other Western funding agencies may also in return for investments in education come to demand measurable and crossnationally comparable evaluations of pupils' performances. At the base is a Western belief in knowledge as essensially quantifiable and objective.

Concluding perspectives

The university is a focal point in the interchange of universal, national and local knowledge. I have here argued for the role of evaluation of the university students in the social construction of the valid knowledge of a culture. By examinations concrete decisions are made concerning the current tension of cross cultural evaluation and evaluation taking in account the values of the local culture.

With the complexities of the issues of educational evaluation of knowledge discussed here, no simple conclusions can be put forth. Some areas of further educational research on examinations, involving continual crossings of disciplinary boundaries, shall be pointed out in conclusion.

On a conceptual level the relation of *knowledge, learning and evaluation* needs to be developed. In contrast to the current often compartmentalised analyses of knowledge, of learning and of evaluation in separate disciplines, the intertwinedness of the three realms need to be adressed.

On an institutional level, a *decoupling of learning and selection* may be considered. Today selection to inequity through grades provides a major motivation for learning. The resulting extrinsic motivation and competion for grades, with high stakes in the outcome, may lead to a simplification of learning and an evaluative erosion of knowledge, contrary to demands of a learning society. A decoupling of learning and selection remains, however, difficult in a meritocratic society which is based on individual achievements as criteria for material privileges.

We may also look for *alternative educational institutions* where evaluation is closely integrated with the production and learning process itself, such as by apprenticeship learning in the crafts and artistry (Kvale, 1993b; Lave & Wenger, 1991). There exists institutions where a unity appears between skills, knowledge and values, such as in the old crafts and in religious orders and in the Asian spiritual and martial arts. In contrast to the public educational system, such institutional niches appear as legitimate to the students who voluntarily enter them. Such alternative educational systems may serve not as models to be imitated, but as inspiration for solutions to some of the problems of evaluation in public education.

When it comes to *crosscultural evaluations* of knowledge, it is necessary to take a closer look at the functions of, and effects of, evaluating knowledge in forms that are commensurable across cultures. Crosscultural interchange of knowledge opens for an enrichment of local and general knowlege. Crosscultural evaluation of knowledge tends to involve a simplification of complex knowledge and trivialisation of controversial values. Current crosscultural evaluations may serve a Western intellectual and economical neocolonisation through education, following the motto - be commensurable or disappear!

References:

Balch, J. (1964). The influence of the evaluating instrument on student's learning. *American Educational Research Journal, 1*, 160-182.

Becker, H., Geer, B. & Hughes, F. (1968). *Making the grade*. New York: Wiley.

Bowles, S. & Gintis, H. (1976). *Schooling in Capitalist America*. London: Routledge and Kegan Paul.

Ch'en, J. (1970) *Mao Papers - anthology and bibliography*. Oxford: Oxford University Press.

Delandshire, G. & Pitrovsky, A.R. (1994). Capturing teachers' knowledge: Performance assessment. *Educational Researher, 23*, No. 5, 11-18.

Dreyfus, H.L. (1993). *What Computers Still Can't Do - A Critique of Artificial Reason*. London: MIT Press.

Durkheim, E. (1977). *The Evolution of Educational Thought*. London: Routledge and Kegan Paul.

Engvik, H., Kvale, S. & Havik, O.E. (1970). Rater reliability in evaluation of essay and oral examinations. *Pedagogisk Forskning - Scandinavian Journal of Educational Research*, 14, 195-220.

Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison*. New York: Viking.

Fredriksen, N. (1984). The real test bias: Influences on teaching and learning. *American Psychologist, 39*, 193-202.

Hansson, F.A. (1993). *Testing testing*. Berkeley: University of California Press.

Jung Chang (1992) *Wild Swans - Three daughters of China*. New York: Flamingo.

Kvale, S. (1970). *En eksaminasjon av universitetseksamener*. Oslo: Universitetsforlaget.

Kvale, S. (1972). *Prüfung und Herrschaft - Hochschulprüfungen zwischen Ritual und Rationalisierung*. Weinheim: Beltz.

Kvale, S. (1980). *Spillet om karakterer i gymnasiet - Interview med elever om bivirkninger av adgangsbegrensning*. København: Munksgaard.

Kvale, S. (Ed.) (1992). *Psychology and Postmodernism*. London: Sage.

Kvale, S. (1990). Evaluation and decentralisation of knowledge. In: M.Granheim, M.Kogan, & U. Lundgren (Eds.) *Evaluation as Policymaking*. London: Jessica Kingsley Publishers, 119-140.

Kvale, S. (Ed.) (1992). *Psychology and Postmodernism*. London: Sage.

Kvale, S. (1993a). Examinations re-examined - Certification of students or of knowledge? In S. Chaiklin & J. Lave (Eds.) *Understanding Practice*. Cambridge: Cambridge University Press, 215-240.

Kvale, S. (1993b) An educational rehabilitation of apprenticeship? Manuscript. Aarhus: University of Aarhus.

Kvale, S. (1995). Validation as communication and action: On the social construction of validity. *Qualitative Inquiry, 1*, no 1, 19-40.

Lave, J. & Wenger, E. (1991). *Situated Learning - Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lyotard, J.F. (1984). *The postmodern condition: A report on knowledge*. Manchester: Manchester University Press.

Madaus, G.F. (1990). Testing as social technology. The Inaugural Boisi Lecture. Boston: Boston College.

Ratio Studiorum et Institutiones Scholastica Societas Jesu. (1887). Berlin: Hoffman.

Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.

Samelson, F. (1987). Was early mental testing: (a) Racist inspired, (b) Objective science, (c) A technology for democracy, (d) The origin of exams, (e) None of the above? (Mark the Right answer). In M.M. Sokal (Ed.) *Psychological Testing and American Society, 1890-1930*. New Brunswick: Rutgers University Press, 113-127.

Thoegersen, S. (1990) *Secondary Education in China after Mao: Reform and Conflict*. Aarhus: Aarhus University Press.

Weiler, H. (1990). Decentralisation in educational governance: An exercise in contradiction. In: M. Granheim, M. Kogan, & U. Lundgren (Eds.) *Evaluation as Policy-making*. London: Jessica Kingsley Publishers, 42-65.

Weiler, H. (1989). Uncertainty and power: New paradigms of knowing and the knowledge base of educational planning. Manuscript. Stanford University.

Wiggins, G.P. (1993). *Assessing student performance*. San Fransisco: Jossey Bass.